IGCP
Instituto de Gestão da Tesouraria
e do Crédito Público, I.P.

Dynamic Models of the Term Structure of Interest Rates

An application of the Nelson and Siegel (1987) approach reformulated by Diebold and Li (2006) and comparison with Principal Component Analysis

Jorge Guedes

July, 2008

Dynamic Models of the Term Structure of Interest Rates

An application of the Nelson and Siegel (1987) approach reformulated by Diebold and Li (2006) and comparison with Principal Component Analysis

Jorge Guedes[1]
IGCP – The Portuguese Treasury and Government Debt Agency
Financial Control Unit
jorge.guedes@igcp.pt

The views expressed in this work are solely the responsibility of the author and should not be interpreted as reflecting the views of IGCP or its members.

[1] I would like to thank Rita Granger from IGCP, for her useful input and interest in this paper. I would also like to thank Paulo Leiria, from IGCP, for his helpful comments. As usual, the author remains entirely responsible for any errors and omissions.

0. Abstract

This study has the purpose of modeling and performing scenario analysis of the Term Structure of Interest Rates (TSIR) with the Nelson and Siegel (N&S) approach. In its essence, this methodology has strong similarities with Principal Components Analysis (PCA), which are illustrated with the results of estimations and simulations presented along the text. In this study, we compare both N&S and PCA in their out-of-sample forecast ability of TSIR. Our approach is to analyze how close the simulated paths of interest rates are to the historical distribution characteristics. Yields, spreads between different maturities and measures of curvature in different parts of TSIR were considered relevant features to be compared. The Nelson and Siegel methodology is reliable to simulate TSIR. With the exception made for skewness and kurtosis, the simulations presented in this work successfully describe the historical distribution of interest rates, spreads and curvatures. This applies to whatever time horizon we choose to make the simulations.

1. Introduction

The term structure of interest rates (TSIR), also often called yield curve, describes the relation between zero coupon interest rates[2] and the corresponding term to maturity. Among other uses, the TSIR provides important information needed in the valuation of financial assets and derivatives, managing financial risk, etc.

The aim of this paper is to assess the predictive ability of different models of TSIR. We are primarily interested in modeling the time series dynamic of TSIR so that we can simulate different random paths that interest rates can follow, and then analyze their probabilistic distribution by computing different descriptive statistics of that distribution. In a government debt management office, this can be useful to compute risk measures such as cost-at-risk[3] or to determine the optimal composition of a benchmark portfolio.

The first difficulty we face when pursuing our target is that TSIR, which is a continuous function of time to maturity, is not directly observable because there are no zero coupon instruments for the whole maturity spectrum.

In reality, even in liquid sovereign debt markets, no more than some points of TSIR are observable. For this reason, if we need an interest rate for a time to maturity for which no instrument is traded, a model of TSIR must be estimated. This need stimulated a huge line of research with the aim of reaching the best possible representation of the entire TSIR from just the few points directly available from market prices.

Despite recent advances in the modeling of TSIR, only in the last decade have efforts been put in forecasting its dynamics or evolution over time. Till then, the concern was primarily to fit the shape of the TSIR, in cross-section, at given moments of time.

---

[2] We identify TSIR with the relationship, at a selected moment in time, between the yields that could be earned in zero-coupon risk-free instruments and the corresponding available maturities. The yield curve concept is more wide-ranging and may refer to coupon or zero coupon rates, and yields may be computed from securities with identical, and not necessarily with the least, credit risk. This definition implies that the yield curve comprises the risk-free zero-coupon curve, or TSIR, as a special case.

[3] Cost-at-risk indicates, with a given probability (usually 95%), the maximum budgetary cost on government debt, for a given period.

Presently, we know that a small number of factors is enough to summarize almost all information contained in the yields of different maturities or the market prices of bills and bonds. For this reason, usually TSIR models are structures composed of a small number of factors and the corresponding weights relating yields of different maturities with those factors.

Out of all of the different methodologies used in the literature, in this study we will only apply the Nelson and Siegel (1987) methodology (N&S) and principal components analysis (PCA), both integrating the suggestion put forward by Diebold and Li (2006).

The first approach, which is very popular among market participants and central banks, is based on fitting the Nelson and Siegel (1987) curve to cross-section data followed by the dynamic modeling of the estimated coefficients' time series, as proposed by Diebold and Li (2006). Actually, this representation is a three-factor dynamic model related with the level, slope and curvature of TSIR.

The N&S approach makes a restriction on factor weights by imposing on them a predefined parametric structure according to the characteristic equation of this methodology, restraining the variety of admissible shapes that TSIR can take. Even though they seem restrictive, conditions are imposed on factor weights with the sole purpose of keeping adherence with fundamental hypotheses such as positive interest rates and discount factors tending to zero with maturity. These are important features in forecasting TSIR, but are not guaranteed in the PCA methodology.

In PCA, factors and the related weights are jointly estimated only imposing the condition that factors are orthogonal (that is, do not show correlation between them) while no provision is enforced on factors' weights. This differs from the N&S approach in which factors may be correlated but weights are predetermined. As a result of PCA, only three factors are often enough to explain between 95 and 99% of the variance of TSIR. Furthermore, the three principal components can be interpreted as level, slope and curvature because they are related to a long-term rate; the spread between short and long-term rates; and a medium-term rate minus an average of short and long-term rates, respectively.

Albeit different in their formulation, estimations and simulations by both methodologies give similar results suggesting that they are closely related. In this study, we compare their out-of-sample forecast ability of TSIR. Our approach is to analyze how close the simulated paths of interest rates are to the historical distribution characteristics. Yields, spreads between different maturities and measures of curvature in different parts of TSIR were considered relevant features to be compared.

This paper is organized in two parts. In the first part, the two methodologies under study will be characterized. The second part presents estimation results and simulations of TSIR with both methodologies.

Part I: Theoretical Analysis

The next section defines the notation used in the rest of the text. In section 3, a non-exhaustive classification of TSIR models in different categories is attempted, with the aim of differentiating PCA and N&S methodologies from alternative approaches of TSIR modeling. Sections 4 and 5 describe in detail PCA and N&S approaches and related estimation methods.

2. Discount Factors, Forward Rates and Spot Rates

Discount factors, forward rates and spot rates are equivalent representations of TSIR so that it is important to define the notation that will be used hereafter.

Let $y_t(\tau)$ be the continuously compounded (zero coupon) spot rate, at moment t , for maturity τ . Then:

- The associated discount factor will be equal to $P_t(\tau) = e^{-y_t(\tau).\tau}$; it starts at 1 for maturity zero and tends asymptotically to zero when maturity tends to infinity.
- The instantaneous forward rate will be equal to $f(\tau) = -\frac{1}{P_t(\tau)} \frac{dP_t(\tau)}{d\tau}$;
- The spot rate, as a function of the forward rate, will be given by $y_t(\tau) = \frac{1}{\tau}\int_0^\tau f_t(u)du$.

3. A non-exhaustive Classification of Models

In modern theory, the two broad classes commonly used to systematize models of TSIR are the equilibrium or fundamental models and no-arbitrage models, also known as reduced-form models. The models considered in this study pertain to another class of models identified as empirical or statistical models. This classification primarily pretends to emphasize the differences between N&S and PCA approaches, and the more traditional models. A more comprehensive systematization of models is beyond the scope of this paper.

3.1. Equilibrium Models

In equilibrium models, TSIR is derived theoretically, assuming consumer's utility maximization and sometimes production functions. The time series dynamics of instantaneous interest rates, given by these models, follow a stochastic differential equation. Interest rates of other maturities are obtained by assuming an affine term structure model and some hypothesis regarding market price of risk. For that reason, these models somewhat disregard the cross-section fit of TSIR and show limited forecast ability. Vasicek (1977) and Cox, Ingersoll and Ross (1985) models pertain to this class of models. Duffie and Kan (1996) provided a multifactor generalization of these models.

3.2. No-Arbitrage Models

Alternatively, no-arbitrage models (such as Hull and White, 1990 and Heath, Jarrow and Morton, 1987) were designed to be perfectly consistent with current TSIR by assuming a functional representation of TSIR that is numerically tractable. No-arbitrage models are estimated under the constraint that the dynamic evolution of interest rates is consistent over time with the shape of TSIR at any given point in time so as to preclude any arbitrage opportunity. While no arbitrage models are flexible modeling tools, they lack economic motivation. On the other hand, while they can attain a good fitting of TSIR to the market data in cross-section, its time series dynamics is rather captured. To obviate this problem some literature suggests doing a recalibration of models frequently.

The remaining models can be grouped in a third class denominated by empirical or statistical models. This category includes, for instance, the spline models based on a piecewise smooth fitting of polynomial functions; the parsimonious models (which include the N&S model) characterized by a functional form with a small number of parameters; and PCA models that are the outcome of the application of this statistical technique to TSIR.

3.3. Empirical Models

McCulloch (1971, 1975) introduced the use of mathematical functions such as polynomials to empirically fit the discount function at given moments of time. After that, several functions were proposed to fit TSIR as a function of discount factors, spot or forward interest rates.

According to the methodology known as “cubic spline”, if maturities spectrum is divided in different ranges, TSIR may be represented by a different third-order polynomial in each of these ranges. Problems may arise in knot points, that is, where polynomials change. In order to guarantee that TSIR is continuous and smooth over all maturities, the polynomial coefficients of two consecutive ranges are estimated subject to the condition that, at knot points, first derivatives must be equal and second derivatives too, simultaneously.

This methodology is flexible enough to fit TSIR with complex shapes. Because it is determined by observations in each group (and the contiguous ones), the more partitions we define the better will be the fit to the observations. Nevertheless, overfitting – as a result of defining too many partitions in the maturity space – may result in inappropriate shapes for TSIR such as those that imply negative forward interest rates. To fix this problem, Fisher et al. (1995, 1996) proposed to fit TSIR directly in forwards rates in place of spot rates or discount factors, and instead of “regression splines” they suggest to use “smoothing splines” so as to preclude overfitting that produces strange shapes in TSIR.

Considering the literature, although it has been successfully used to fit TSIR in cross-section, this methodology is not used to forecast or simulate the future evolution of TSIR. There seem to be some reasons behind this fact. On the one hand, the definition of knot points at each moment of time is always a subjective task. On the other, it is not straightforward to interpret the estimated polynomial coefficients and to model them in time series is even more complicated. These difficulties together make it unfeasible to use cubic splines to forecast or simulate interest rates in an effective way.

The most important shortcoming of spline models is the great number of parameters to be estimated. This is a considerable limitation because market participants and central banks prefer to fit parsimonious functions on the number of parameters and easy to implement. According to BIS (2005), nine of the thirteen central banks covered by their study have been using the Nelson and Siegel methodology or the Svensson (1994) extension, which are functions of TSIR with only four and five coefficients, respectively.

Recently, this kind of models has gained a new impulse with Diebold and Li (2006), who interpret the coefficients of Nelson and Siegel function as latent factors (alike those obtained by PCA) and model them as autoregressive processes obtaining promising results when they forecast TSIR one step ahead. In the same line of research, Diebold, Rudebusch and Aruoba (2005) and Diebold, Piazzesi and Rudebusch (2005) incorporate macroeconomic variables related with factors that would help in forecasting.

4. Principal Components Analysis

Principal components analysis (PCA) is a statistical technique to decompose a set of correlated variables into factors by means of identifying common behavior patterns. Through an analysis of relations between variables in the system, PCA tries to reproduce them with a smaller number of variables, by taking advantage of redundant information. In other words, PCA intends to reduce the problem dimensionality, with little loss of information, making it more comprehensible.

The basic idea of PCA is to generate a new set of variables that are linear combinations of the former set and account for most of the variance of the original set. The transformed variables are called principal components or factors and are orthogonal between them. This means that the only non-null cells of the covariance matrix are the diagonal variances, with obvious computational advantages.

Provided that there is a strong relationship between interest rates for different maturities, PCA can simplify TSIR by making use of only a small number of independent factors. Stock and Watson (1988), and Litterman and Scheinkman (1991) were the pioneers in applying PCA to the study of TSIR. Using United States data, their work identifies three factors accountable for 98% of interest rate variance. Those factors are responsible for movements in the level, steepness and curvature of TSIR.

4.1. Methodology characterization

Interest rates for different maturities are highly correlated. This means that there are only some few independent sources of information shared by all variables in the multivariate system called TSIR. PCA is a statistical method for extracting uncorrelated sources of variation, making it possible to represent the information of a $n \times k$ matrix as a linear combination of k uncorrelated variables, called factors or principal components. It also computes the proportion of variance of the original matrix that is explained by each principal component.

The purpose of PCA is to reduce dimensionality so that only the most relevant sources of information are considered through the first m principal components, those that explain the most important part of variance of TSIR. Furthermore, the principal components are orthogonal, thus their covariance matrix is diagonal: we only have to compute m variances instead of $k(k+1)/2$ elements of the $k \times k$ matrix of variances and covariances of the original system.

The advantage of PCA is that the system will be a function of only m variables instead of k, which simplifies tremendously calculations and sources of uncertainty, with little loss of accuracy.

PCA is based on the analysis of eigenvalues and eigenvectors of $\mathrm{V} = \mathrm{X'X/T}$, a $k \times k$ symmetric matrix of correlations between variables in X. Each principal component is a linear combination of columns in X, the weights $(w_{1m}, w_{2m}, \ldots, w_{km})$ being selected so that:

- The first principal component explains the greatest part of total variance in X; the second principal component explains the greatest part of the remaining variance, and so on.
- The principal components are uncorrelated.

This can be attained selecting the weights from the eigenvectors of the correlation matrix.

Denoting by W the $k \times k$ eigenvectors matrix of V, we have:

$\mathrm{VW} = \mathrm{W}\Lambda$,

where Λ is $k \times k$ diagonal matrix of eigenvalues of V. If $\mathrm{W} = (w_{i,j})$ for $i, j = 1, \ldots, k$, then the m-th column of W, represented by $w_m = (w_{1m}, w_{2m}, \ldots, w_{km})$ is the $k \times 1$ eigenvector associated to the eigenvalue λ_m.

Once the eigenvectors of the correlation matrix have been obtained, the next step is to sort them according to the eigenvalues, from the greatest to smallest. This gives us the principal components by their significance order.

If we sort the columns of W ascendingly by the associated eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_k$ then the m-th principal component is given by:

$P_m = w_{1m}X_1 + w_{2m}X_2 + \ldots + w_{km}X_k = Xw_m$, where X_i represents the i-th column of X, i.e., the historical data of the standardized variable i. Each principal component is a linear combination of the variables represented in the system. In matrix notation, the $T \times m$ matrix of principal components which has P_m in the m-th column, may be represent by:

$$P = XW.$$

Owing to the sorting of columns of W, principal components thus obtained are sorted too, so that P_1 is associated to eigenvalue λ_1, P_2 to λ_2, etc. According to PCA methodology, the eigenvalues are equal to the proportion of variance explained by each factor. The variance of the first principal component is equal to the maximum eigenvalue, the variance of the second principal component is equal to the second highest eigenvalue and so on and so forth. Thus, the proportion of total variance that is explained by the m-th principal component is equal to $\lambda_m \Big/ \sum_{i=1}^{k} \lambda_i$. However, because in the case of standardized variables, the sum of eigenvalues is equal to the number of variables, the proportion of variance explained by the n first principal components is equal to $\sum_{i=1}^{n} \lambda_i / k$.

As $W' = W^{-1}$, then $X = PW'$, that is,

$$X_i = w_{i1}P_1 + w_{i2}P_2 + \ldots + w_{ik}P_k.$$

Each vector of data may be described as a linear combination of principal components. All k principal components would be necessary to explain entirely the total variance of the system. It is customary to consider only the most significant principal components because they are sufficient to explain the most important part of that variance with little loss of information that will be the lesser, the smaller the eigenvalues associated to the excluded principal components.

4.2. Some Considerations on Data Standardization and Stationarity

Due to the fact that PCA is not scale invariant, the data to be used should be standardized (that is, have zero mean and unit variance) so that principal components are not dominated by the variable with the greatest variance[4]. Therefore, PCA should be implemented with variables subtracted from their mean and divided by their standard deviation.

Another important consideration relates to the consequences of non-stationarity of data for the computation of principal components. According to Machado et al. (2001), although this technique was initially developed under the assumption of stationarity, PCA is still applicable with non-stationary variables. With regard to the standardization, it is adequate when the original variables are stationary. However, when variables are integrated of order one and some of the series exhibit strong increasing or decreasing trends, the common standardization is no longer recommended because variance is no longer a good measure of volatility.

---

[4] It is customary to previously standardize the original series in order to get comparable data if the variances of variables differ much, or if the units of measurement of the variables differ. Otherwise, we actually can have PCA implemented in both ways: if we proceed

The question of whether interest rates can be characterized as mean-reverting or unit root processes, that is, as stationary or non-stationary processes, is controversial in literature. If interest rates contain unit roots, then shocks in interest rates are permanent and their effects will persist in the course of time. On the other hand, if interest rates can be characterized as stationary processes, then innovations are transitory and are expected to be reversed after a while. Following this reasoning, and given that interest rates are bounded by zero and should not tend to infinity, we should be expecting that empirical studies would lead to the conclusion that interest rates are stationary.

Nevertheless, the first studies on unit-root tests reported that the unit root hypothesis could not be rejected and, as a result, the assumption that interest rates are non-stationary has, in fact, been taken as established in the following research. More recently, some doubts were raised concerning the power of unit roots tests[5] under certain circumstances. According to the latter evidence, the failure to reject unit-root hypotheses seems to be more the result of lack of power of standard tests rather than being evidence against stationarity of interest rates. For this reason, a stationarity analysis of nominal interest rates is beyond the scope of this paper. Nevertheless, even if interest rates should be considered as non-stationary, this should not prevent us from applying the usual standardization because the existence of differences in trends does not seem relevant in interest rates for different maturities.

4.3. Monte Carlo Simulation using PCA

The basic idea is to apply Monte Carlo simulation to generate the future distribution of TSIR according to the PCA representation. To achieve that goal, we must model the time series behavior of factors and different methodologies can be used. We will only refer to random walk (RW), autoregressive (AR) and vector autoregressive (VAR) processes.

As shown in PCA methodology characterization, by definition, we have:

$\mathrm{P} = \mathrm{XW}$,

where P represents the principal components, X stands for the original data standardized and W is the eigenvectors matrix.

Since the eigenvectors matrix is orthogonal, by multiplying both sides of the late equation by $\mathrm{W'}$, we get:

$\mathrm{PW'} = \mathrm{XWW'} = \mathrm{X}$ or $\mathrm{X} = \mathrm{PW'}$.

By letting $x_{1t}, x_{2t}, ..., x_{nt}$ stand for the changes in standardized daily interest rates for different n maturities and letting $p_{1t}, p_{2t}, ..., p_{3t}$ stand for the three factors, we attain the following matrix notation to represent the changes in interest rates:

$x_{it} = w'_{1i}\, p_{1t} + w'_{2i}\, p_{2t} + w'_{3i}\, p_{3t} + \xi_{it}$ for $i = 1, ..., n$ where ξ_{it} is the error motivated by only considering three factors in the model and w'_{1i}, w'_{2i} and w'_{3i} are the sensibilities of each interest rate to the factors.

---

without standardize we will be doing a PCA of the covariance matrix, while a PCA in standardized variables is a PCA done on the correlation matrix.

[5] The most popular unit root tests are the augmented Dickley and Fuller (ADF) and Phillips and Perron (PP) tests.

We consider that the first three principal components are the risk factors and the rest of the variance of interest rates is a noise.

In order to model the dynamics of x_{it}, the time series properties of factors (p_{1t}, p_{2t}, p_{3t}) must be analyzed. We will consider three different models: a random walk, a first-order autoregressive process, $AR(1)$, which can be demonstrated to exhibit mean reversion, and a first-order vector autoregressive process, $VAR(1)$, that also considers the interrelations between factors.

Once we have estimated the models, it is enough to generate random numbers according to the normal distribution and simulate the three factors (p_{1t}, p_{2t}, p_{3t}) dynamics. In the case of PCA, because correlation between p_{1t}, p_{2t}, p_{3t} is equal to zero, by construction, the simulation can be independent, i.e., $AR(1)$ and $VAR(1)$ should return the same results, if errors are independent.

It should be noticed that the interest rates resulting from simulation are standardized. For that reason, to return to the original vector, standardized interest rates must be multiplied by the standard deviation and added to the mean.

5. The Nelson and Siegel approach

Nelson and Siegel (1987) (N&S) proposed to fit TSIR using a smooth and flexible function capable of representing, with a small number of parameters – the reason why we included it in the parsimonious models class – a large variety of typical shapes that TSIR can take in the course of time. Even knowing that the Nelson and Siegel function has more empirical than theoretical motivation, its widespread use by market players and central banks is irrefutable proof of its popularity.

Several extensions to N&S, such as Svensson (1994) and Bliss (1996) have been proposed, in order to add in more flexibility in the estimation and to adequately fit more complicated shapes of TSIR. It is obvious that models with more factors cannot fit worse than less flexible models, however they impose a burden in terms of further difficulties in the estimation.

The model parameters, assumed to be time-changing, are estimated for each cross-section sample at equal time intervals depending on the frequency of observations. In the literature, frequency of observations usually ranges from daily to monthly. Since variations in parameters are related to changes in the shape of TSIR, following the results of Diebold and Li (2006), if we apply time series estimation methods to model and simulate the parameters, we will also be able to forecast the future evolution of TSIR.

5.1. Nelson and Siegel (1987)

In the parametric model proposed by Nelson and Siegel, the relationship between interest rates and maturity is derived from the assumption that spot rates follow a 2nd order differential equation and that forward rates, being forecasts of future spot rates, are the solution of that equation with two identical real roots.

The instantaneous forward rates, in moment t and maturity τ are represented by the following exponential expansion:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\tau/\lambda_t} + \beta_{3,t}\, e^{-\tau/\lambda_t}\tau/\lambda_t \,.$$

Integrating from 0 to τ, we get the spot rates function:

$$y_t(\tau) = \frac{1}{\tau}\int_0^{\tau} f_t(m)dm = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right) + \beta_{3,t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right)$$

where τ stands for maturity and $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$ and λ_t stand for the four parameters to be estimated.

This framework has the adequate features to capture the usual shapes of TSIR. One of them is that limits exist for interest rates of very long and short maturities. In spot as well as forward rates function, the parameters have the following interpretations. For long maturities, interest rates tend asymptotically to $\beta_{1,t}$ (i.e., the very long-term interest rate), $\beta_{1,t} + \beta_{2,t}$ is the initial value of the curve (i.e., the instantaneous interest rate) and the spread between long and short-term interest rates is equal to $-\beta_{2,t}$ (i.e. the average slope). Parameters $\beta_{3,t}$ and λ_t have no direct economic interpretation but determine the way the transition between the short and long ends of the curve is made and are responsible for the hump that TSIR presents. The magnitude of the hump is determined by the $\beta_{3,t}$ value while the orientation of the hump is determined by its sign: if $\beta_{3,t} < 0$ then the hump will present a "U" shape and if $\beta_{3,t} > 0$ it will look like an inverted "U". If $\beta_{3,t}$ is approximately zero, instead of a maximum or minimum, TSIR will be a monotone increasing or decreasing function depending on whether $\beta_{2,t}$ is negative or positive, respectively.

The velocity at which interest rates move toward their long-term value is determined by the decay parameter λ_t, which is equal to the maturity for which the weight associated to factor $\beta_{3,t}$ in the forward rates function is maximized. Furthermore, the forward rate has an extreme value for $\tau = \left(1 - \frac{\beta_{2,t}}{\beta_{3,t}}\right)\lambda_t$. The smaller (greater) λ_t is, the smaller (greater) will be the maturity for which forward rates reach their extreme value and the faster (later) they will converge to the long-term value.

Since interest rates must be strictly positive, the model will make sense only if the following conditions are imposed:

$\beta_{1,t} > 0$

$\beta_{1,t} + \beta_{2,t} > 0$

$\lambda_t > 0$

The first two inequalities guarantee that both short and long-tem limits of TSIR are positive while the third inequality is in accordance with the function having a long-term asymptote.

5.2. Bliss (1996) and Svensson (1994) Extensions to Nelson and Siegel

Despite the fact that the Nelson and Siegel model is able to reproduce a great variety of TSIR shapes, more flexible mathematical functions have been proposed in the literature with the aim of fitting even better to the shapes observed in the market. We only detail the extensions to Nelson and Siegel proposed by Bliss (1996) and Svensson (1994).

The improvement proposed by Bliss (1996) is to allow components associated to steepness and curvature to depend on different decay parameters: λ_{1t} and λ_{2t}. The corresponding forward rate function is:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\tau/\lambda_{1t}} + \beta_{3,t}\, e^{-\tau/\lambda_{2t}} \tau/\lambda_{2t} ,$$

so that spot rates curve will be represented by:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\tau/\lambda_{1t}}}{\tau/\lambda_{1t}}\right) + \beta_{3,t}\left(\frac{1-e^{-\tau/\lambda_{2t}}}{\tau/\lambda_{2t}} - e^{-\tau/\lambda_{2t}}\right).$$

In case $\lambda_{1t} = \lambda_{2t}$, Bliss curve reduces to the Nelson and Siegel one.

Svensson (1994) suggests the inclusion of an additional factor with an independent decay parameter to bring in a second hump that would result in a better fitting to existing TSIR. The resulting forward rates curve is:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\tau/\lambda_{1t}} + \beta_{3,t}\, e^{-\tau/\lambda_{1t}} \tau/\lambda_{1t} + \beta_{4,t}\, e^{-\tau/\lambda_{2t}} \tau/\lambda_{2t} ,$$

and the associated spot rates curve is:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\tau/\lambda_{1t}}}{\tau/\lambda_{1t}}\right) + \beta_{3,t}\left(\frac{1-e^{-\tau/\lambda_{1t}}}{\tau/\lambda_{1t}} - e^{-\tau/\lambda_{1t}}\right) + \beta_{4,t}\left(\frac{1-e^{-\tau/\lambda_{2t}}}{\tau/\lambda_{2t}} - e^{-\tau/\lambda_{2t}}\right).$$

When $\beta_{4t} = 0$, the Svensson curve reduces to the Nelson and Siegel one.

When $\lambda_{1t} = \lambda_{2t}$, the Svensson curve reduces to Nelson and Siegel but with a factor associated to curvature equal to $\beta_{3,t} + \beta_{4,t}$, i.e., the curvature effect is distributed through two parameters. When estimating, if we arrive at $\lambda_{1t} \approx \lambda_{2t}$, a multicollinearity problem arises.

5.3. Nelson and Siegel as a 3-Factor Dynamical Model

Originally, the Nelson and Siegel model was designed to fit the entire TSIR in a static approach, cross-sectionally, at selected moments of time. Lately, Diebold and Li (2006) adapted the exponential components structure of Nelson and Siegel to model TSIR as a model with 3 parameters evolving dynamically in time. Diebold and Li show that Nelson and Siegel, reinterpreted in light of a dynamic model, despite its simple structure, is able to represent the empirical properties of historically observed TSIR and obtain encouraging results when interest rates forecasts are compared with those of alternative models. These outcomes opened the way to the study of the time series behavior of the factors as well as their relationship with relevant macroeconomic variables. These efforts added an economic content to the Nelson and Siegel approach, which it was lacking initially.

In the original Nelson and Siegel formula, the parameter λ_t changes in time. Nevertheless, according to Diebold and Li (2006), there is no loss of generality if we assume that λ_t is constant, $\lambda_t = \lambda$ because this parameter only determines the maturity for which the weight associated to the curvature factor is maximized and has no obvious economic interpretation. When we fix λ, the characteristic equation of Nelson and Siegel is linear in parameters and so it can be simply estimated by ordinary least squares. In empirical studies, there are three different ways to fix λ. The simplest one presets it at an arbitrary value usually ranging from 2 to 3 years. Another approach is to set it equal to the average of the time changing λ_t that were estimated cross-sectionally over the whole sample period. In the third method, the one we chose to follow in this study, the constant λ is estimated jointly with the other time-changing parameters, within the whole sample, by non-linear least squares.

The three components of the model, i.e. $\left(1, \frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t}, \frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right)$ may be interpreted as the weights of each factor. Representing them as a function of maturity, some results are worth mentioning. The weight associated to the first factor is a constant and does not depend on maturity. The weight associated to coefficient $\beta_{2,t}$ is a decreasing function of τ. The weight of the third coefficient starts at zero for $\tau = 0$, increases to intermediate values of τ, and decreases to zero for long maturities presenting the shape of a concave function of τ. These results are identical to those obtained with PCA and validate the interpretation of coefficients $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$ as latent factors related with level, steepness and curvature, respectively. Graphically, we can perceive that an increase of $\beta_{1,t}$ raises interest rates equally over all maturities; an increase in $\beta_{2,t}$ raises short-term more than long-term rates and as a result of an increase of $\beta_{3,t}$, intermediate maturities rates are increased more than short and long-term rates.

The following state-space representation can embrace all dynamic models of the Nelson and Siegel type:

$y_t = X_t \beta_t + \varepsilon_t$ (measure equation)

$$\begin{bmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-\tau_1\lambda}}{\tau_1\lambda} & \frac{1-e^{-\tau_1\lambda}}{\tau_1\lambda} - e^{-\tau_1\lambda} \\ 1 & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} - e^{-\tau_2\lambda} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} - e^{-\tau_2\lambda} \end{bmatrix} \begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{bmatrix}$$

$\beta_t = \mu + \Phi \beta_{t-1} + \upsilon_t$ (state or forecast equation)

In case of $VAR(1)$:

$$\begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \upsilon_{2,t} \\ \upsilon_{3,t} \end{bmatrix}$$

In case of $AR(1)$:

$$\begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \phi_{11} & 0 & 0 \\ 0 & \phi_{22} & 0 \\ 0 & 0 & \phi_{33} \end{bmatrix} \begin{bmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \upsilon_{2,t} \\ \upsilon_{3,t} \end{bmatrix}$$

where $\begin{bmatrix} \varepsilon_t \\ \upsilon_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0_{N\times 1} \\ 0_{K\times 1} \end{bmatrix}, \begin{bmatrix} H & 0 \\ 0 & Q \end{bmatrix} \right)$

The measure equation defines the $(T \times N)$ vector of actual interest rates as the sum of factors multiplied by their weights as stated by the model and a vector of normal errors that are independent across maturities. The error variance for each maturity $\sigma_t^2(\tau_i)$ with $i = 1,\ldots,N$ compose a $(N \times N)$ diagonal matrix H. The $(K \times 1)$ vector β_t stands for the factors and X_t correspond to the $(N \times K)$ matrix of factors' weights that will be time changing only if the decay parameter λ_t is variable.

The vector μ has dimensions $(K \times 1)$ and the matrix Φ has dimension $(K \times K)$, and will be diagonal or complete according to whether we specify an $AR(1)$ or $VAR(1)$ model, in that order. In the case of a random walk, $\mu = 0$ and $\Phi = 1$. The errors in the state equation follow a normal distribution with a $(K \times K)$ covariance matrix Q, which we will assume to be complete.

The assumption that matrix H is diagonal implies that deviations from TSIR are not correlated across maturities. This is common in literature and decreases the number of parameters to be estimated.

The assumption of an unrestricted matrix Q allows that shocks on the three factors (betas) are correlated although according to the PCA methodology, we do not expect any correlation between factors.

5.4. Estimation Methods

5.4.1 Two-Stage Estimation

In a first stage, the measure equations are estimated in cross-section samples with a regular frequency. Parameter estimators are obtained for each moment of time. In a second stage, the time series dynamics of parameters are specified and estimated as an $AR(1)$, $VAR(1)$ or random walk process.

If we assume that the decay parameter is constant, the Nelson and Siegel equation becomes linear and it can be estimated, in the first stage, by ordinary least squares. However, the selection of λ value is not obvious and will always entail some degree of subjectivity.

The alternative is to estimate the decay parameter jointly with other parameters but, in that case, the estimation must make use of non-linear least squares or maximum likelihood methods. As usual in literature, in the second stage, the decay parameters dynamics is not modeled explicitly and we must make use of the constant estimator if we had estimated the system imposing a constant λ or the average of time varying λ_t, if we let it change across time.

In the case that λ is constant, and a non-linear least squares method is used, the objective function to be minimized is the sum of squared deviations between observed and estimated interest rates:

$$Min_{\{\beta_1,\beta_2,\beta_3,\lambda\}}\sum_{t=1}^{T}\left[y_t(\tau)-\hat{y}_t(\tau)\right]^2 = Min_{\{\beta_1,\beta_2,\beta_3,\lambda\}}\sum_{t=1}^{T}\left[y_t(\tau)-\beta_{1t}-\beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)-\beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}-e^{-\lambda\tau}\right)\right]^2$$

The previous equation can be estimated using the fminsearch functionality of Matlab® that uses the simplex method of Nelder and Mead (1965). This is a direct search numerical method for minimizing an objective function in a many-dimensional space, using no information from derivatives, which do not exist for this problem.

5.4.2 Kalman Filter

An alternative to the two-stage estimation is to estimate jointly all model parameters in the state-space formulation by maximum likelihood through a Kalman filter. The simultaneous estimation of all parameters in only one stage has advantages in terms of statistical inference over the two-stage estimation. To see this we should notice that the second-stage parameter estimation does not consider the errors of the first stage while the joint estimation of measure and state equations does.

The likelihood of the system given by measure and forecast equations is a function of the set of parameters $\Theta=(\lambda_1,\lambda_2,\beta_t,\mu,\Phi,H,Q)$:

$$L=\sum_{t=1}^{T}\left[-\frac{1}{2}\ln(2\pi)-\frac{1}{2}\ln\left(\left|f_{t|t-1}\right|\right)-\frac{1}{2}\eta'_{t|t-1}\,f_{t|t-1}^{-1}\eta_{t|t-1}\right]$$

where:

$\eta_{t|t-1} = y_t - y_{t|t-1}$ stands for the $(N \times 1)$ vector of forecast errors;

$y_{t|t-1}$ stands for the vector of interest rates forecasts given the information available at the moment $t-1$,

$f_{t|t-1} = E(\eta_{t|t-1}\eta'_{t|t-1})$ corresponds to the covariance matrix of forecast errors.

The parameters are estimated by maximum likelihood:

$$Max\left[-\frac{T}{2}\ln(2\pi\sigma_\varepsilon) - \frac{1}{2}\sum_{j=1}^{T}\left(\frac{Y_j - Y_j(\beta)}{\sigma_\varepsilon}\right)^2\right]$$

where $\hat{\sigma}_\varepsilon = \sqrt{\frac{1}{T}\sum_{j=1}^{T}\left[Y_j - Y_j(\beta)\right]^2}$.

The most important danger in the Nelson and Siegel estimation through Kalman Filter is to obtain estimators that are local and not global optima, i.e. the risk of false convergence. The optimization algorithms usually used – non-linear least squares and maximum likelihood – are solved based on initial values. The greater the sensibility of parameter estimators to the initial values used, the greater should be the caution to be employed. For this reason, it is convenient to use a grid of initial values and not only one. If the solution does not converge always to the same value, some scrutiny should be made when selecting the optimum.

Nevertheless, the interpretation of parameters β_1 and β_2 given before suggests that we should use as initial values the longest maturity interest rate and the difference between this and the shortest maturity interest rate, respectively. Since for β_3 and λ there is no specific economic interpretation, several initial values should be used. In the case of parameter λ, because it must be positive, it is recommended using a grid of values ranging from 1 to the longest maturity in the data. The bounds of initial values for parameter β_3 should be set based on the results of some preliminary estimations[6].

Another practical rule is to use the coefficients estimated at a given date as initial values of the subsequent observation in the sample.

[6] Using this method, Geyer and Mader (1999) allow initial values of β_3 to vary between –1 and +1.

Part II: Empirical Analysis

This part of the work begins with a description of the data used in the empirical analysis. In section 7, the results of Nelson and Siegel and PCA estimations are presented. Equipped with the estimated coefficients, TSIR is simulated according to both models. Section 8 analyses the probabilistic distribution of interest rates, spreads between maturities and curvature measured along different points of TSIR. The two models are compared with the characteristics observed in the historical sample. Section 10 concludes this study giving some directions for future research.

6. Data

The inputs to be used in this work are the EURIBOR (Euro Interbank Offered Rate) and swap rates. Given that we are interested in modeling risk-free rates we should be using instead Treasury bills and bonds. Alternatively, the former can be converted in the latest (and vice-versa) through an appropriate modeling of credit spreads, which we will not pursue here. EURIBOR rates are already zero-coupon rates; swap rates will be adequately[7] converted to zero-coupon rates.

There are two ways of proceeding with the empirical analysis. One is to use yields; the other is to convert yields in prices and use them in estimations. Nevertheless, minimization of the sum of squared errors in yields has benefits over prices. Since bond prices for short maturities are insensitive to interest rates ($P_{t,\tau} = e^{-r\tau}$ tends to 1, when τ tends to zero independently of interest rate r) and ordinary least squares give more weight to the greatest errors, minimization of errors in prices would give less importance to short-term errors than minimization of errors in interest rates. On the other hand, both methodologies perform equally well in the long-term.

Daily EURIBOR for maturities between 1 month and 1 year, and Euro Swap Rates between 2 years and 10 years, 15, 20, 25 and 30 years were obtained from Reuters®. The data set is comprised of 25 time series, with 2212 observations, covering the time period from 04/01/1999 to 23/08/2007. In a preliminary treatment of data, market rates were converted into continuously compounded rates given market conventions, value dates and considered maturities. Effective terms to maturity for each interest rate were also computed and stored in the database. Figure 1 illustrates the TSIR historical data used in this study.

[7] See Galitz (1995), page 167, equations 9.9 and 9.10.

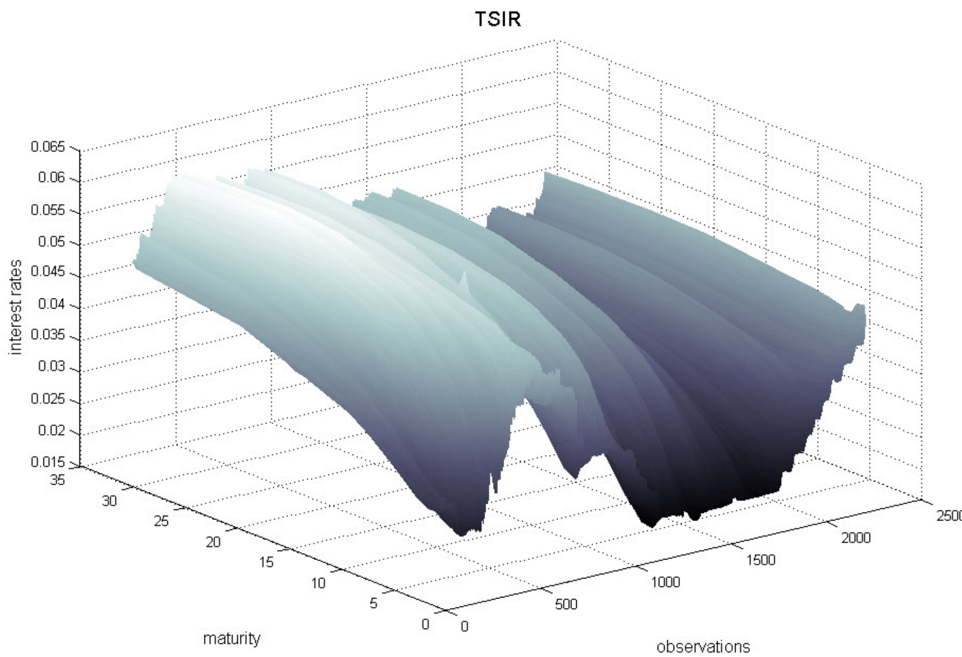

Figure 1: Term Structures of Interest Rates in the period from 04/01/1999 to 23/08/2007

Figure 2 depicts the average, 5th and 95th percentiles, standard deviation and first order autocorrelation for interest rates (upper panel) and the same statistics for their first differences or variations (lower panel).

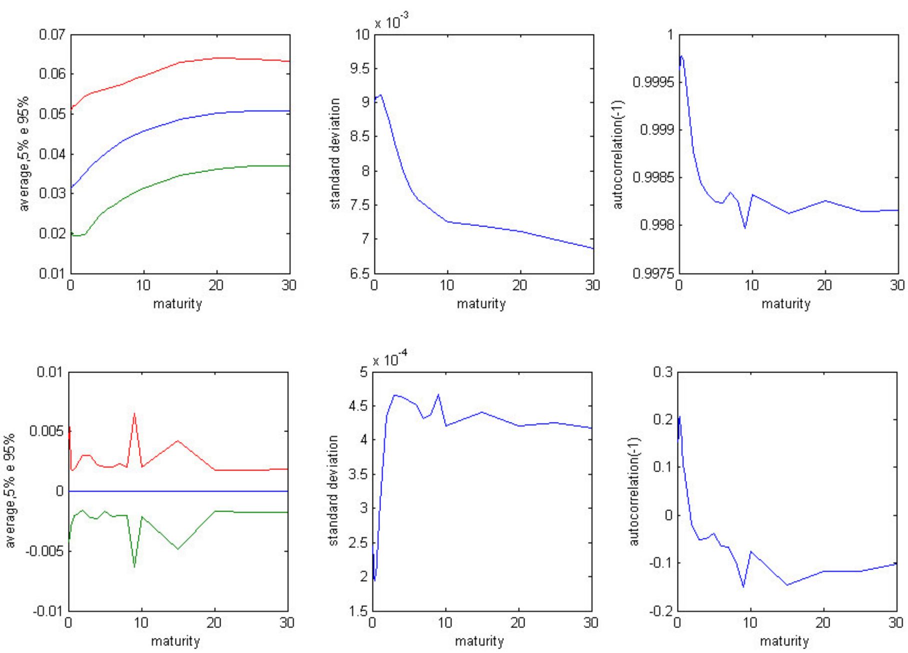

Figure 2: Interest Rates and Interest Rates Variations

The average TSIR is upward sloping and concave. Interest rate volatility decreases with maturity and interest rates are rather persistent contrary to their first differences that seem stationary.

7. Estimation Results

In this study, λ (constant) was estimated jointly with the other (time changing) parameters by non-linear least squares, starting from a battery of different initial values. All of them converged to the same solution: 1,8825 years.

In figure 3, we present the sum of squared errors of Nelson and Siegel and 3-factor PCA estimations, for the period covered by the sample. When the whole sample is considered, the PCA methodology is at least as good as Nelson and Siegel, and for the last 500 observations (last 2 years), PCA shows to fit better than the Nelson and Siegel function.

Figure 4 illustrates for given dates, the effective and estimated TSIR by N&S and PCA. In the first four observations, there are no differences between the two methods but in the last two, the PCA fit is visibly better.

When taken together, figures 3 and 4 reveal some difficulty of N&S to capture the shape of TSIR in the most recent period, recommending the use of some extension of the N&S model. The Bliss and Svensson functions reach a better fit but with the burden of overparametrization. The algorithms become dependent on initial values and may not converge to a unique solution.

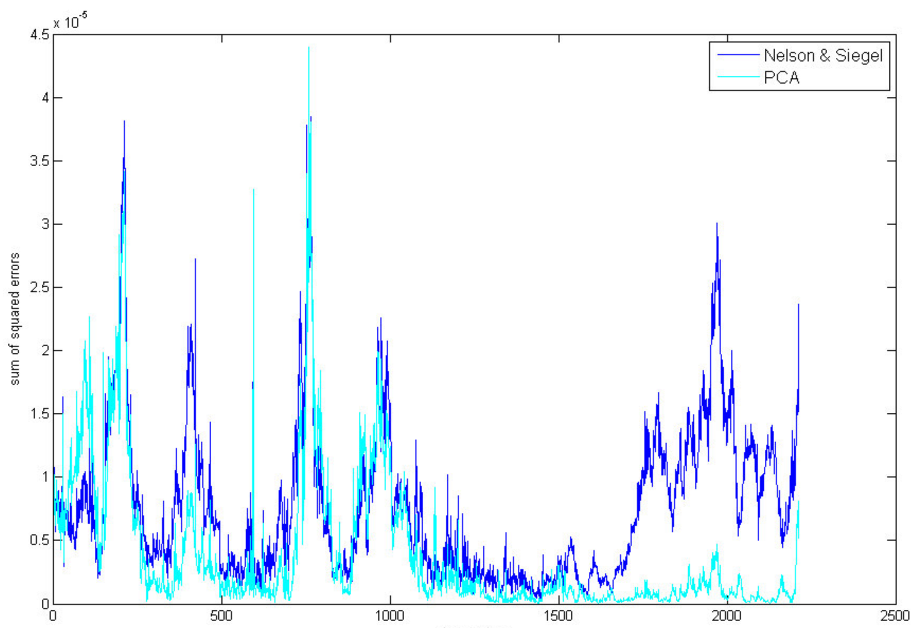

Figure 3: Sum of squared errors of N&S and PCA

The preference for Nelson and Siegel (in detriment of Bliss and Svensson functions) prevails in literature and is a consequence of the fact that, from a practical implementation perspective, ordinary least squares is a more straightforward estimation method and does not suffer from convergence and initial values problems which non-linear least squares or maximum likelihood estimations may reveal. On the other hand, the additional explaining ability of Bliss and Svensson will always be marginal. Finally, a simpler but more intuitive model will be more valuable than a more complex and better-fitting model that lacks economic interpretation or that is unrelated to economic variables.

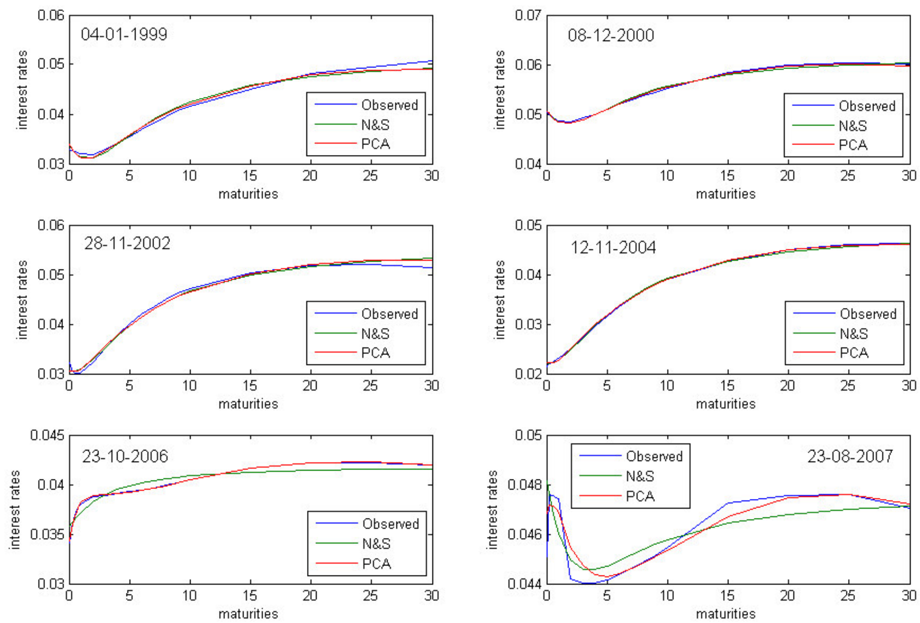

Figure 4: Effective and Estimated TSIR by N&S and PCA at given dates

		Average	St. Dev.	Skewness	Kurtosis	Min	Max	Correlation Matrix			AR(-1)	AR(-20)	AR(-250)
Betas	1	0,05337	0,00764	-0,42154	1,92400	0,03743	0,06651	1,000	-0,451	-0,270	0,998	0,977	0,782
	2	-0,02174	0,00962	0,29614	2,12970	-0,03827	0,00082		1,000	0,491	0,998	0,971	0,276
	3	-0,01930	0,01526	-0,05083	1,72700	-0,05517	0,00814			1,000	0,996	0,922	-0,172
Factors	1	0,00000	4,61760	0,32666	2,00680	-7,61120	9,64240	1,000	0,000	0,000	0,999	0,983	0,541
	2	0,00000	1,79540	-0,28702	2,32930	-4,08710	3,78740		1,000	0,000	0,997	0,960	0,347
	3	0,00000	0,63900	0,52234	2,21180	-1,32510	1,62510			1,000	0,994	0,898	-0,438
Δ Betas	1	-0,00000	0,00045	0,18884	4,49690	-0,00205	0,00252	1,000	-0,897	-0,041	-0,128	-0,014	-0,008
	2	0,00000	0,00058	-0,34664	6,11960	-0,00384	0,00350		1,000	-0,188	-0,155	-0,008	-0,026
	3	0,00001	0,00135	0,43863	5,72430	-0,00529	0,00685			1,000	0,018	-0,013	0,004
Δ Factors	1	0,00259	0,15199	0,29763	4,53510	-0,79995	0,70595	1,000	0,668	-0,604	0,120	-0,014	-0,009
	2	-0,00131	0,13598	0,33473	4,70770	-0,65357	0,75876		1,000	-0,474	-0,161	-0,013	-0,024
	3	-0,00023	0,06713	-0,61693	6,44750	-0,43160	0,26100			1,000	-0,026	-0,010	-0,006

Table 1: Descriptive Statistics of N&S Betas and PCA Factors

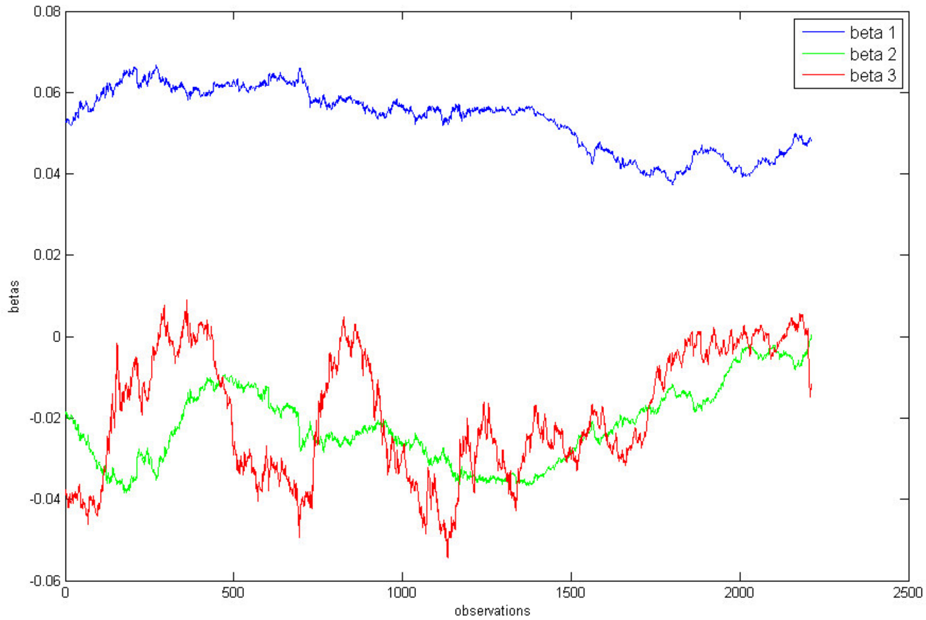

Figure 5: N&S Betas

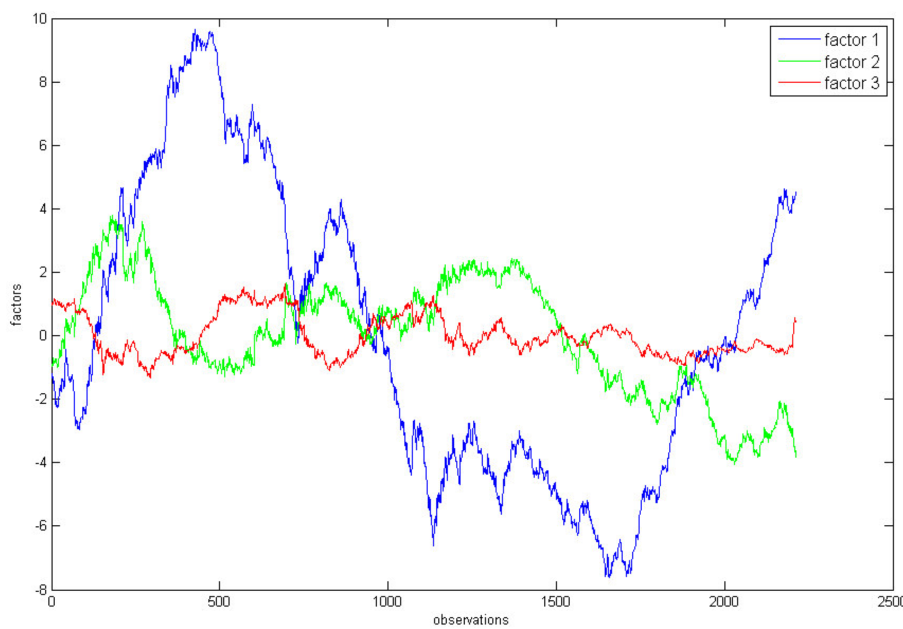

Figure 6: PCA Factors (Principal Components)

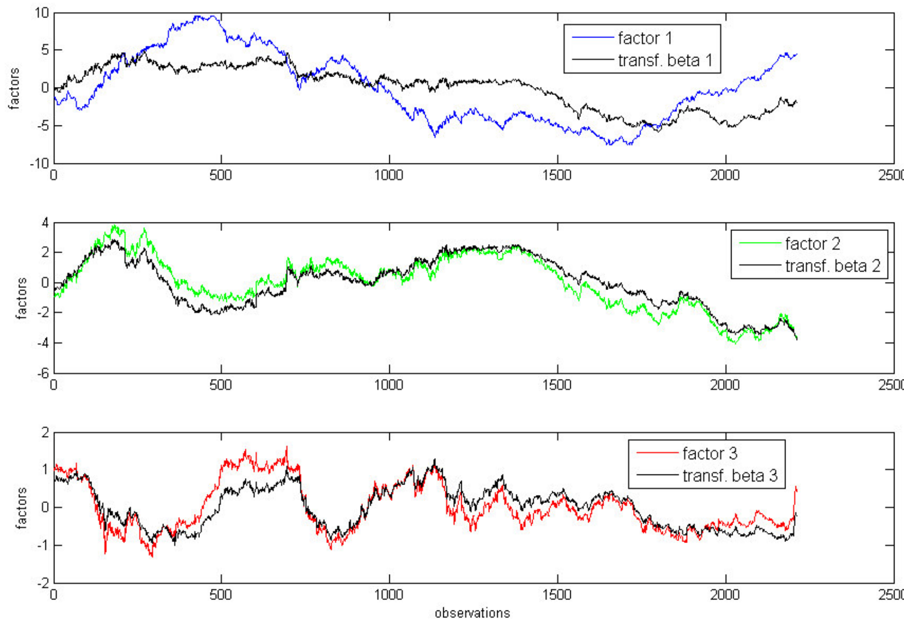

Figure 7: PCA factors and linear transformations of N&S Betas

In figure 5, the beta coefficients of the Nelson and Siegel function are presented for the whole period under study. The beta 2 coefficient which is equal to the difference between the short and the long-term is always negative with the exception of the final period. TSIR is positively sloped but the last 3 years are marked by a continuous reduction of spreads between short and long-terms. In the face of beta 3 behavior, we can say that curvature was almost always negative and when it was positive it showed small values, that is, TSIR was almost always concave and sometimes flat.

Figure 6 displays the three PCA factors or principal components. Although they have different scales or levels, they summarize the same information of N&S betas, as is illustrated in figure 7. When betas are linearly transformed (that is, multiplied by a scalar and a constant is added to the result), the dynamics of betas is comparable to that of factors.

By inspecting the weights associated to betas and factors, as displayed in figure 8, we reach the conclusion that the information conveyed by both methodologies is the same.

The first factor and beta 1 correspond to parallel displacements of TSIR. Shocks on that factor generate variations of equal sign and amplitude over all TSIR. The second factor and beta 2 are responsible for changes in TSIR gradient because shocks on this factor cause displacements of opposite sign in the ends of the curve: short and long-terms. Finally, shocks on the third factor or beta 3 have effects in the ends of the curve of opposite sign of the effect originated on the medium-term of the curve.

The first factor or beta is associated to the level of interest rates; the second and the third to the steepness and curvature, respectively, albeit with different signs in the two methodologies.

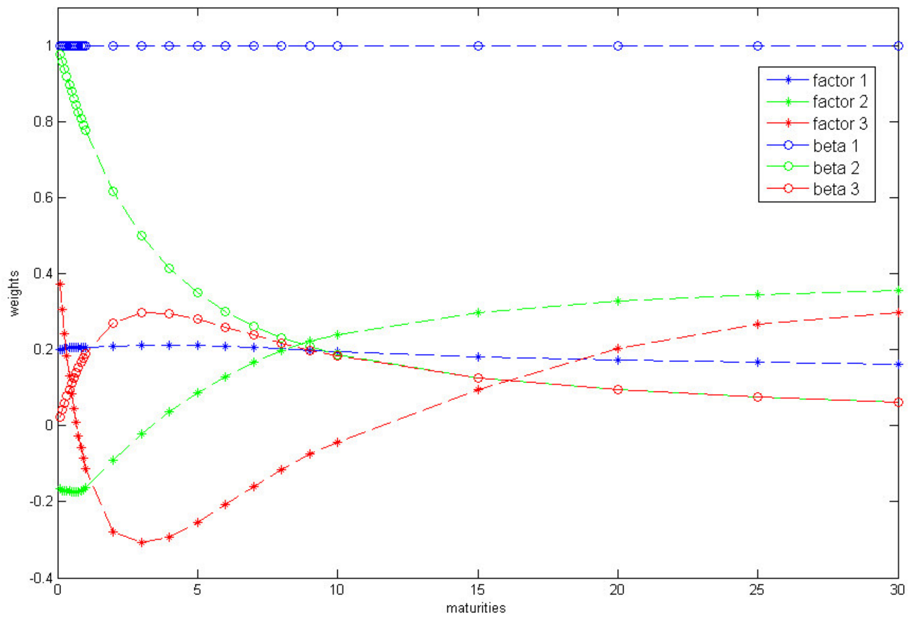

Figure 9: PCA factor weights and N&S beta weights

After having estimated in a first stage the factors in PCA and betas in N&S, in the second stage they are modeled in time series. Table 2 displays the outcomes of estimation of $AR(1)$ and $VAR(1)$ models, which will be useful in the simulation of TSIR dynamics according to both methodologies.

Considering the small correlation between factors, the $VAR(1)$ model adds little to the explaining power of an $AR(1)$ model, as expected. However, we will use the former because it is slightly richer in information.

8. Simulation Outcomes

In tables 3, 4 and 5, we present some descriptive statistics of the distribution of interest rates for different maturities, of spreads between interest rates of several maturities and some measures of curvature obtained across TSIR span. The selected statistics are: average, standard deviation, skewness and kurtosis coefficients, Jarque-Bera normality test and corresponding probability, minimum and maximum values and 5 and 95-th percentiles.

We compare the historical distribution (the upper panel of the tables) with the distribution of simulated rates at the end of 1 and 5 years (middle and lower panels).

There are no significant differences in the simulations made through both methods: TSIR is typically upward sloping and concave and short-term interest rates are more volatile than long-term ones.

The histograms of interest rates, spreads and curvatures simulated by the N&S methodology are presented in figures 9 (interest rates), 10 (spreads) and 11 (curvatures). As there are no visible differences regarding the outcomes of the PCA methodology, we do not display the related histograms.

At the end of one year, the historical and simulated distributions have similar expected values and standard deviations. The primary difference resides in the shape (skewness and kurtosis) of the distribution: while normality is rejected in historical interest rates, spreads and curvatures, the same is not true for the simulated distributions. One of the main assumptions of simulations is that interest rates follow a normal distribution, which is far from occurring in reality because they typically show positive skewness (they are bounded by the zero value) and excess of kurtosis. Actually, of 2500 generated simulations, it was necessary to eliminate 19 of the N&S method and 21 of the PCA method, because they took negative values at some moment in time (in the following 5 years).

This fact suggests that, if it is intended to enhance the models, the issue that should be given more attention is the non-normality of interest rates, in order to match them with empirical distributions.

Nelson and Siegel

		VAR(1)			AR(1)		
		Beta 1	Beta 2	Beta 3	Beta 1	Beta 2	Beta 3
Coefficients	Beta 1	0,99922	0,00029	-0,00517	0,99838		
	Beta 2	0,00229	0,99523	-0,00676		0,99936	
	Beta 3	-0,00086	0,00543	0,99717			0,99577
	Const.	0,00007	-0,00001	0,00009	0,00008	-0,00001	-0,00007
t-statistics	Beta 1	708,9	0,2	-1,2	794,8		
	Beta 2	1,9	622,5	-1,8		770,2	
	Beta 3	-1,2	5,8	459,2			527,3
	Const.	1,1	-0,1	0,4	1,2	-0,2	-1,5
Probabilities	Beta 1	0,00	0,87	0,22	0,00		
	Beta 2	0,06	0,00	0,07		0,00	
	Beta 3	0,23	0,00	0,00			0,00
	Const.	0,29	0,95	0,68	0,21	0,86	0,13
R^2		0,99652	0,99635	0,99213	0,99652	0,99629	0,99212

Principal Component Analysis

		VAR(1)			AR(1)		
		Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3
Coefficients	Factor 1	0,99966	-0,00081	0,00078	0,99966		
	Factor 2	-0,00157	0,99811	-0,00112		0,99811	
	Factor 3	-0,02201	0,01853	0,99407			0,99407
	Const.	0,00259	-0,00131	-0,00023	0,00260	-0,00131	-0,00024
t-statistics	Factor 1	1433,2	-1,3	2,5	1427,5		
	Factor 2	-0,9	621,4	-1,4		619,1	
	Factor 3	-4,4	4,1	446,1			445,5
	Const.	0,8	-0,5	-0,2	0,8	-0,5	-0,2
Probabilities	Factor 1	0,00	0,20	0,01	0,00		
	Factor 2	0,38	0,00	0,16		0,00	
	Factor 3	0,00	0,00	0,00			0,00
	Const.	0,42	0,65	0,87	0,42	0,65	0,87
R^2		0,99893	0,99432	0,98903	0,99892	0,99427	0,98899

Table 2: Estimation Results of $AR(1)$ and $VAR(1)$ models

	Nelson and Siegel											Principal Components Analysis										
	Maturity	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)	Maturity	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)
Historical	'1M'	3,132%	0,90%	0,52	2,13	170,0	0,00%	2,042%	5,105%	2,101%	4,861%	'1M'	3,132%	0,90%	0,52	2,13	170,0	0,00%	2,042%	5,105%	2,101%	4,861%
	'1Y'	3,305%	0,91%	0,35	1,97	142,4	0,00%	1,937%	5,274%	2,157%	4,980%	'1Y'	3,305%	0,91%	0,35	1,97	142,4	0,00%	1,937%	5,274%	2,157%	4,980%
	'5Y'	4,023%	0,77%	0,24	2,01	111,4	0,00%	2,596%	5,629%	2,858%	5,396%	'5Y'	4,023%	0,77%	0,24	2,01	111,4	0,00%	2,596%	5,629%	2,858%	5,396%
	'10Y'	4,560%	0,72%	0,13	1,98	102,8	0,00%	3,122%	5,936%	3,408%	5,744%	'10Y'	4,560%	0,72%	0,13	1,98	102,8	0,00%	3,122%	5,936%	3,408%	5,744%
	'15Y'	4,852%	0,72%	0,03	1,94	104,3	0,00%	3,428%	6,288%	3,687%	5,944%	'15Y'	4,852%	0,72%	0,03	1,94	104,3	0,00%	3,428%	6,288%	3,687%	5,944%
	'30Y'	5,048%	0,68%	-0,17	1,93	116,9	0,00%	3,695%	6,313%	3,928%	6,013%	'30Y'	5,048%	0,68%	-0,17	1,93	116,9	0,00%	3,695%	6,313%	3,928%	6,013%
At the end of 1 year	'1M'	3,793%	0,70%	-0,03	2,97	0,4	83,39%	1,346%	5,980%	2,620%	4,951%	'1M'	3,806%	0,65%	-0,02	2,87	1,9	37,93%	1,779%	6,219%	2,707%	4,878%
	'1Y'	3,691%	0,72%	-0,06	3,03	1,3	51,09%	1,260%	6,175%	2,476%	4,871%	'1Y'	3,710%	0,74%	-0,05	2,80	5,6	6,01%	1,384%	6,190%	2,456%	4,938%
	'5Y'	4,192%	0,64%	-0,08	3,04	3,1	21,27%	1,889%	6,617%	3,113%	5,225%	'5Y'	4,176%	0,62%	-0,09	2,79	8,1	1,76%	2,191%	6,129%	3,150%	5,164%
	'10Y'	4,749%	0,53%	-0,05	2,99	0,9	62,95%	2,849%	6,680%	3,867%	5,604%	'10Y'	4,709%	0,53%	-0,09	2,89	5,0	8,20%	2,954%	6,420%	3,835%	5,553%
	'15Y'	5,018%	0,49%	0,00	2,96	0,2	88,86%	3,201%	6,688%	4,192%	5,830%	'15Y'	5,017%	0,50%	-0,09	2,93	4,2	12,51%	3,415%	6,668%	4,193%	5,812%
	'30Y'	5,301%	0,49%	0,04	2,93	1,2	56,13%	3,579%	6,846%	4,495%	6,137%	'30Y'	5,240%	0,45%	-0,09	2,96	3,9	14,57%	3,780%	6,689%	4,465%	5,977%
At the end of 5 years	'1M'	3,442%	0,96%	-0,07	2,99	2,2	33,14%	0,079%	6,475%	1,825%	4,994%	'1M'	3,403%	0,94%	-0,08	2,79	7,2	2,74%	0,281%	6,478%	1,814%	4,945%
	'1Y'	3,512%	0,90%	-0,04	2,94	1,2	53,71%	0,578%	6,469%	2,003%	4,985%	'1Y'	3,529%	0,92%	-0,09	2,81	7,2	2,77%	0,166%	6,348%	2,005%	5,036%
	'5Y'	4,189%	0,76%	-0,01	2,90	1,1	58,12%	1,848%	6,799%	2,948%	5,415%	'5Y'	4,180%	0,75%	-0,08	2,99	2,6	27,21%	0,947%	6,739%	2,915%	5,421%
	'10Y'	4,719%	0,69%	-0,03	2,96	0,7	72,07%	2,439%	7,139%	3,551%	5,805%	'10Y'	4,699%	0,71%	-0,05	3,09	2,0	35,96%	1,977%	7,147%	3,502%	5,906%
	'15Y'	4,963%	0,69%	-0,04	2,98	0,7	70,18%	2,632%	7,356%	3,812%	6,054%	'15Y'	4,986%	0,71%	-0,04	3,12	2,0	37,51%	2,555%	7,462%	3,800%	6,181%
	'30Y'	5,218%	0,71%	-0,04	3,01	0,7	71,73%	2,786%	7,579%	4,030%	6,367%	'30Y'	5,172%	0,68%	-0,01	3,13	1,7	42,77%	2,802%	7,503%	4,042%	6,292%

Table 3: Simulation outcomes of N&S and PCA – descriptive statistics of interest rates

Nelson and Siegel

	Spread	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)
Historical	5Y-1Y	0,718%	0,42%	-0,13	1,98	101,8	0,00%	-0,344%	1,594%	0,016%	1,335%
	10Y-1Y	1,255%	0,62%	-0,37	2,07	131,1	0,00%	-0,209%	2,335%	0,110%	2,109%
	15Y-1Y	1,547%	0,72%	-0,39	2,12	128,4	0,00%	-0,048%	2,776%	0,207%	2,523%
	5Y-1M	0,891%	0,53%	0,23	2,55	38,4	0,00%	-0,254%	2,382%	0,024%	1,798%
	10Y-1M	1,428%	0,65%	0,09	2,27	52,3	0,00%	0,025%	3,008%	0,398%	2,431%
	15Y-1M	1,720%	0,72%	0,01	2,24	53,7	0,00%	0,186%	3,394%	0,566%	2,779%
At the end of 1 year	5Y-1Y	0,500%	0,33%	-0,02	2,97	101,8	0,00%	-0,759%	1,506%	-0,043%	1,029%
	10Y-1Y	1,058%	0,51%	-0,07	2,96	131,1	0,00%	-0,745%	2,695%	0,196%	1,884%
	15Y-1Y	1,326%	0,61%	-0,07	2,97	128,4	0,00%	-0,693%	3,311%	0,301%	2,314%
	5Y-1M	0,399%	0,49%	0,04	3,00	38,4	0,00%	-1,402%	2,002%	-0,396%	1,215%
	10Y-1M	0,957%	0,61%	-0,02	2,96	52,3	0,00%	-1,388%	2,827%	-0,061%	1,941%
	15Y-1M	1,225%	0,68%	-0,04	2,96	53,7	0,00%	-1,336%	3,268%	0,108%	2,326%
At the end of 5 years	5Y-1Y	0,676%	0,41%	0,03	2,93	0,8	66,47%	-0,606%	2,211%	0,007%	1,346%
	10Y-1Y	1,207%	0,62%	0,05	2,95	0,9	63,02%	-0,791%	3,330%	0,198%	2,221%
	15Y-1Y	1,451%	0,73%	0,06	2,96	1,0	59,73%	-0,978%	3,832%	0,260%	2,642%
	5Y-1M	0,746%	0,61%	0,01	2,94	1,1	56,63%	-1,257%	3,009%	-0,253%	1,733%
	10Y-1M	1,277%	0,76%	0,03	2,93	1,2	53,75%	-1,115%	4,128%	0,027%	2,526%
	15Y-1M	1,520%	0,84%	0,04	2,93	1,3	51,09%	-1,120%	4,595%	0,152%	2,882%

Principal Components Analysis

	Spread	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)
Historical	5Y-1Y	0,718%	0,42%	-0,13	1,98	101,8	0,00%	-0,344%	1,594%	0,016%	1,335%
	10Y-1Y	1,255%	0,62%	-0,37	2,07	131,1	0,00%	-0,209%	2,335%	0,110%	2,109%
	15Y-1Y	1,547%	0,72%	-0,39	2,12	128,4	0,00%	-0,048%	2,776%	0,207%	2,523%
	5Y-1M	0,891%	0,53%	0,23	2,55	38,4	0,00%	-0,254%	2,382%	0,024%	1,798%
	10Y-1M	1,428%	0,65%	0,09	2,27	52,3	0,00%	0,025%	3,008%	0,398%	2,431%
	15Y-1M	1,720%	0,72%	0,01	2,24	53,7	0,00%	0,186%	3,394%	0,566%	2,779%
At the end of 1 year	5Y-1Y	0,466%	0,35%	-0,10	2,98	101,8	0,00%	-0,916%	1,580%	-0,126%	1,043%
	10Y-1Y	0,999%	0,54%	-0,08	2,93	131,1	0,00%	-1,036%	2,614%	0,098%	1,881%
	15Y-1Y	1,307%	0,62%	-0,07	2,92	128,4	0,00%	-0,985%	3,117%	0,267%	2,320%
	5Y-1M	0,371%	0,45%	-0,11	3,09	38,4	0,00%	-1,253%	1,747%	-0,362%	1,101%
	10Y-1M	0,904%	0,56%	-0,11	3,04	52,3	0,00%	-1,228%	2,687%	-0,032%	1,821%
	15Y-1M	1,211%	0,61%	-0,10	3,01	53,7	0,00%	-1,177%	3,191%	0,184%	2,234%
At the end of 5 years	5Y-1Y	0,650%	0,45%	0,01	2,91	4,3	11,55%	-0,863%	2,171%	-0,089%	1,401%
	10Y-1Y	1,170%	0,67%	0,02	2,91	4,6	10,08%	-1,260%	3,239%	0,080%	2,296%
	15Y-1Y	1,457%	0,76%	0,03	2,90	4,9	8,74%	-1,367%	3,861%	0,241%	2,753%
	5Y-1M	0,777%	0,59%	-0,02	2,91	5,1	7,72%	-1,061%	2,967%	-0,203%	1,723%
	10Y-1M	1,297%	0,73%	0,00	2,90	5,4	6,80%	-0,998%	3,925%	0,083%	2,516%
	15Y-1M	1,583%	0,80%	0,01	2,90	5,6	6,01%	-0,940%	4,337%	0,283%	2,904%

Table 4: Simulation outcomes of N&S and PCA – descriptive statistics of spreads

Nelson and Siegel

	Curvature	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)
Historical	2*5Y-(1Y+10Y)	0,182%	0,28%	0,27	2,07	106,1	0,00%	-0,523%	0,870%	-0,209%	0,649%
	2*10Y-(5Y+15Y)	0,245%	0,14%	-0,41	2,07	142,7	0,00%	-0,160%	0,760%	-0,007%	0,431%
	2*15Y-(10Y+20Y)	0,137%	0,07%	-1,05	18,63	22859,0	0,00%	-0,731%	0,488%	0,041%	0,234%
	2*5Y-(1M+10Y)	0,355%	0,50%	-0,01	2,51	22,3	0,00%	-0,803%	1,763%	-0,487%	1,161%
	2*10Y-(1M+20Y)	0,981%	0,60%	0,22	2,50	41,5	0,00%	-0,264%	2,593%	0,013%	2,031%
	2*15Y-(1M+30Y)	1,524%	0,68%	0,21	2,48	40,9	0,00%	0,195%	3,356%	0,485%	2,579%
At the end of 5 years	2*5Y-(1Y+10Y)	-0,058%	0,25%	0,06	3,05	1,8	41,41%	-0,840%	0,832%	-0,461%	0,353%
	2*10Y-(5Y+15Y)	0,290%	0,12%	-0,07	2,97	2,4	30,73%	-0,127%	0,692%	0,081%	0,492%
	2*15Y-(10Y+20Y)	0,127%	0,05%	-0,06	2,97	1,5	47,50%	-0,059%	0,298%	0,037%	0,214%
	2*5Y-(1M+10Y)	-0,159%	0,47%	0,06	3,05	1,7	42,58%	-1,617%	1,477%	-0,906%	0,601%
	2*10Y-(1M+20Y)	0,548%	0,53%	0,03	2,99	0,3	85,37%	-1,470%	2,263%	-0,334%	1,440%
	2*15Y-(1M+30Y)	0,941%	0,61%	-0,02	2,97	0,3	87,50%	-1,398%	2,809%	-0,069%	1,923%
At the end of 5 years	2*5Y-(1Y+10Y)	0,146%	0,30%	-0,03	2,99	0,5	77,38%	-0,979%	1,092%	-0,351%	0,637%
	2*10Y-(5Y+15Y)	0,287%	0,14%	0,06	3,00	1,5	48,28%	-0,207%	0,788%	0,058%	0,516%
	2*15Y-(10Y+20Y)	0,116%	0,06%	0,06	3,01	1,5	47,45%	-0,088%	0,337%	0,019%	0,212%
	2*5Y-(1M+10Y)	0,215%	0,55%	-0,04	3,00	0,6	73,11%	-1,847%	1,890%	-0,692%	1,113%
	2*10Y-(1M+20Y)	0,906%	0,67%	0,01	2,94	0,5	76,59%	-1,186%	3,420%	-0,203%	1,994%
	2*15Y-(1M+30Y)	1,265%	0,76%	0,03	2,93	1,1	58,75%	-1,120%	4,113%	0,018%	2,511%

Principal Components Analysis

	Curvature	Average	Standard Deviation	Skewness	Kurtiosis	JB	Prob	Min	Max	Perc.(5%)	Perc.(95%)
Historical	2*5Y-(1Y+10Y)	0,182%	0,28%	0,27	2,07	106,1	0,00%	-0,523%	0,870%	-0,209%	0,649%
	2*10Y-(5Y+15Y)	0,245%	0,14%	-0,41	2,07	142,7	0,00%	-0,160%	0,760%	-0,007%	0,431%
	2*15Y-(10Y+20Y)	0,137%	0,07%	-1,05	18,63	22859,0	0,00%	-0,731%	0,488%	0,041%	0,234%
	2*5Y-(1M+10Y)	0,355%	0,50%	-0,01	2,51	22,3	0,00%	-0,803%	1,763%	-0,487%	1,161%
	2*10Y-(1M+20Y)	0,981%	0,60%	0,22	2,50	41,5	0,00%	-0,264%	2,593%	0,013%	2,031%
	2*15Y-(1M+30Y)	1,524%	0,68%	0,21	2,48	40,9	0,00%	0,195%	3,356%	0,485%	2,579%
At the end of 5 years	2*5Y-(1Y+10Y)	-0,067%	0,23%	-0,11	3,08	5,7	5,80%	-0,882%	0,622%	-0,436%	0,304%
	2*10Y-(5Y+15Y)	0,225%	0,12%	-0,04	2,90	2,0	37,38%	-0,171%	0,584%	0,034%	0,415%
	2*15Y-(10Y+20Y)	0,133%	0,04%	-0,06	2,89	2,6	27,85%	-0,005%	0,260%	0,067%	0,199%
	2*5Y-(1M+10Y)	-0,162%	0,43%	-0,09	3,02	3,2	20,51%	-1,718%	1,175%	-0,878%	0,533%
	2*10Y-(1M+20Y)	0,422%	0,51%	-0,11	3,08	5,7	5,87%	-1,439%	1,957%	-0,412%	1,252%
	2*15Y-(1M+30Y)	0,988%	0,58%	-0,12	3,04	5,7	5,79%	-1,204%	2,787%	0,016%	1,933%
At the end of 5 years	2*5Y-(1Y+10Y)	0,131%	0,29%	-0,03	2,93	1,0	59,45%	-0,804%	1,213%	-0,361%	0,604%
	2*10Y-(5Y+15Y)	0,233%	0,14%	0,03	2,92	1,1	57,91%	-0,290%	0,667%	0,011%	0,463%
	2*15Y-(10Y+20Y)	0,134%	0,05%	0,02	2,95	0,5	76,82%	-0,051%	0,279%	0,056%	0,217%
	2*5Y-(1M+10Y)	0,257%	0,53%	-0,05	2,93	1,7	41,97%	-1,560%	2,103%	-0,625%	1,097%
	2*10Y-(1M+20Y)	0,858%	0,67%	-0,02	2,91	0,9	62,25%	-1,260%	3,341%	-0,259%	1,931%
	2*15Y-(1M+30Y)	1,398%	0,75%	-0,01	2,92	0,7	69,58%	-0,998%	4,136%	0,155%	2,641%

Table 5: Simulation outcomes of N&S and PCA – descriptive statistics of curvatures

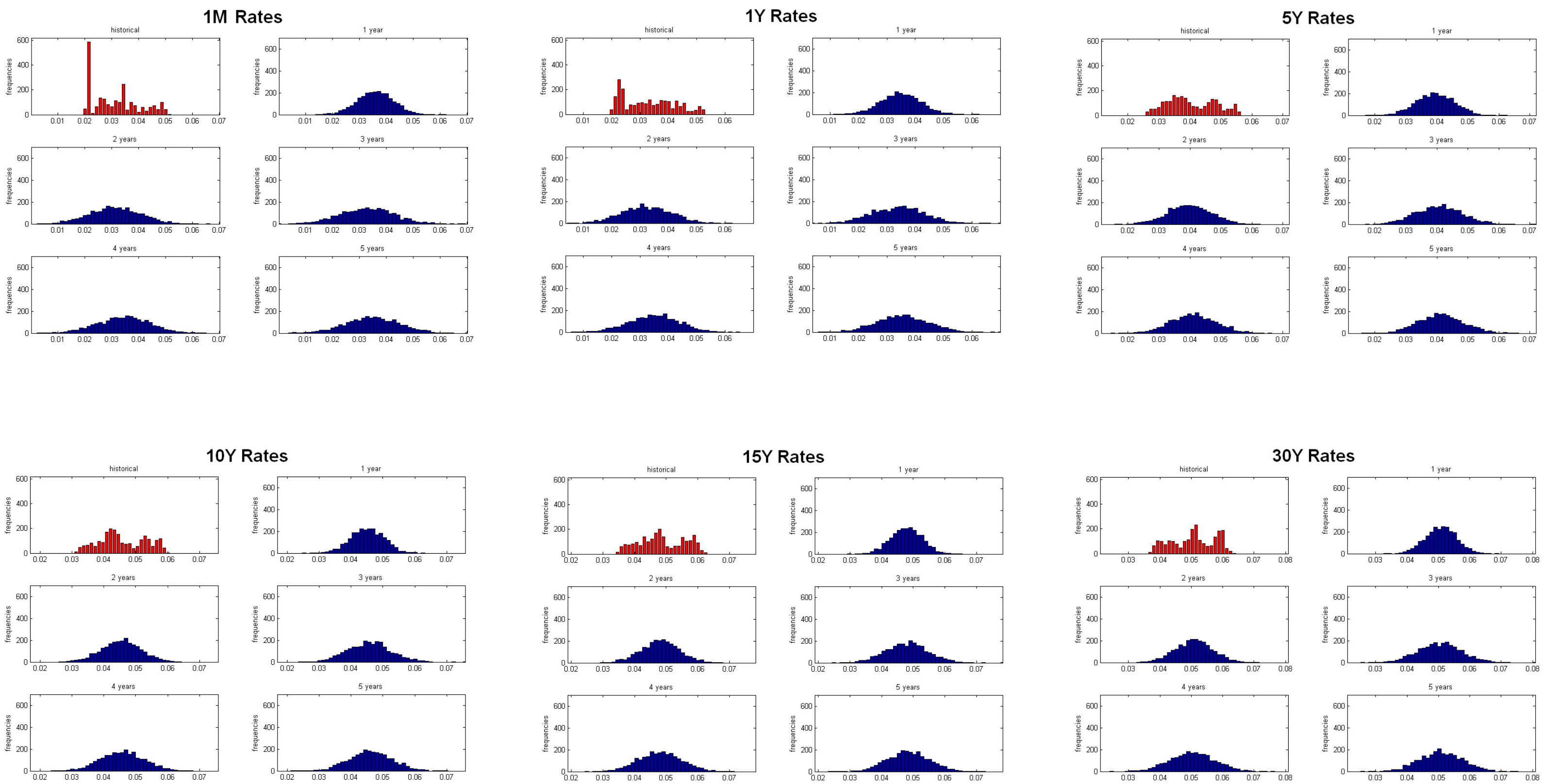

Figure 9: Simulation outcomes of N&S – histograms of interest rates

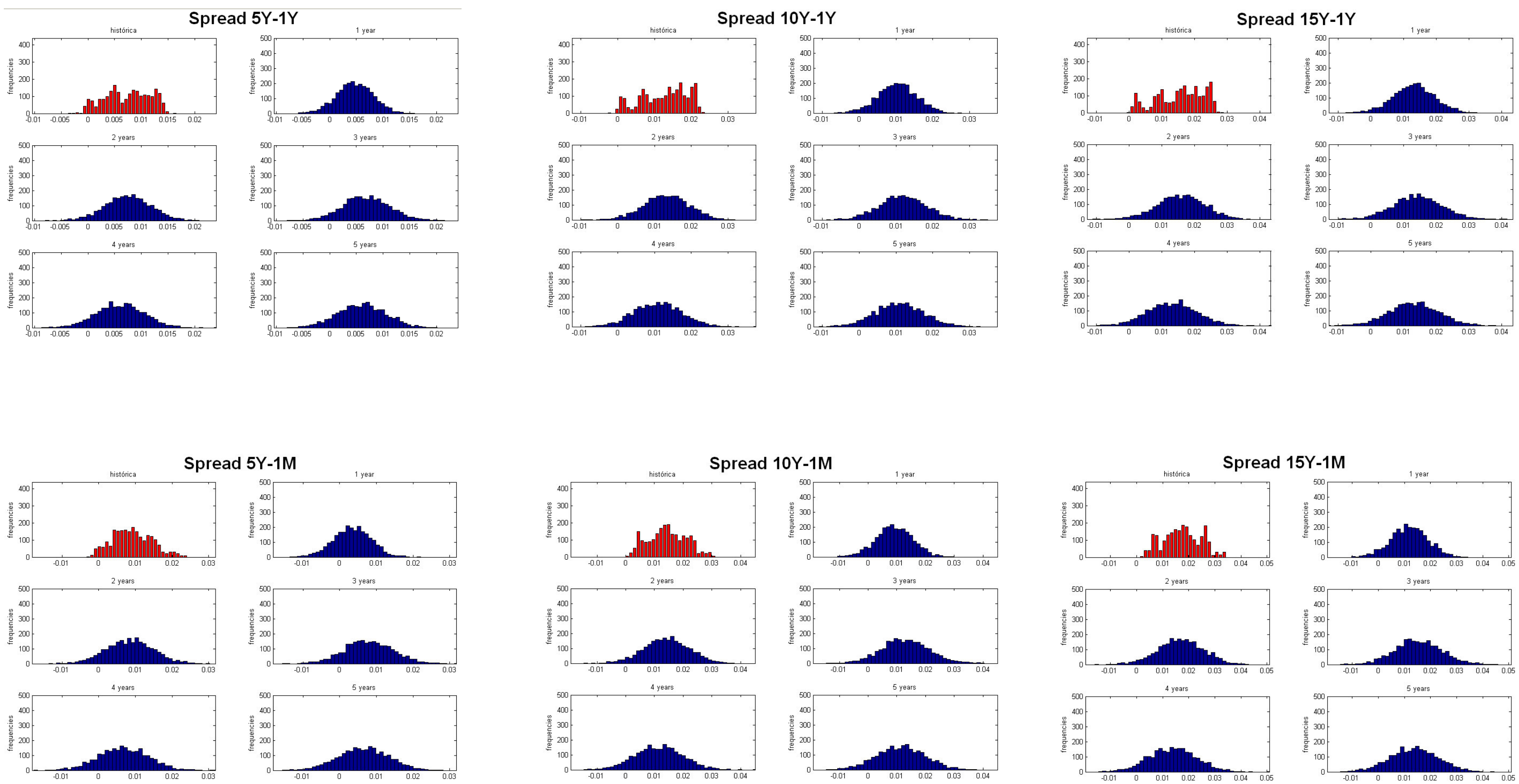

Figure 10: Simulation outcomes of N&S – histograms of spreads

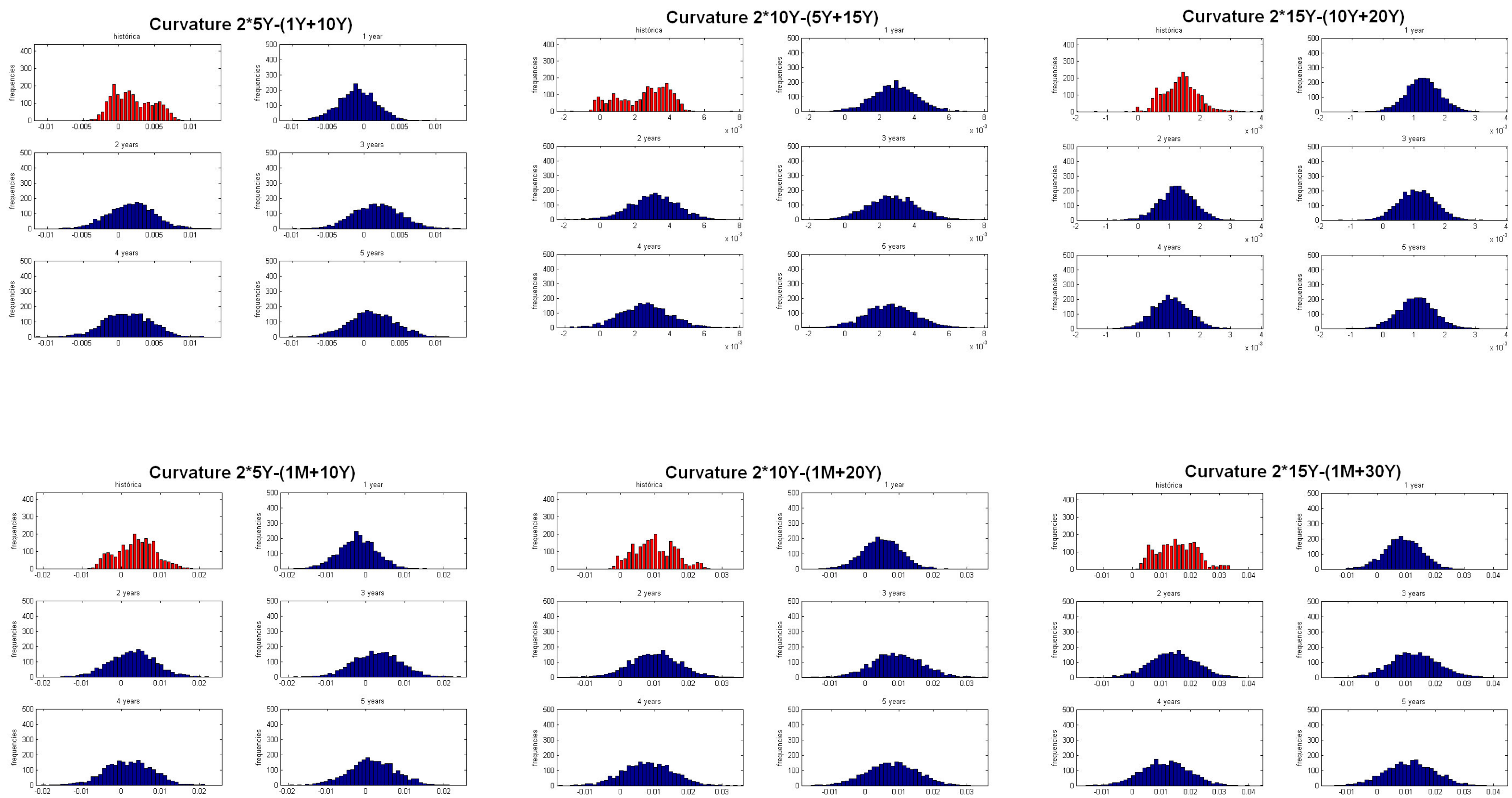

Figure 11: Simulation outcomes of N&S – histograms of curvatures

9. Conclusion

This study had the purpose of modeling and performing scenario analysis of the TSIR with the N&S approach.

In its essence, this methodology has strong similarities with PCA, which are illustrated with the results of estimations and simulations presented along the text. PCA would have some gains over N&S in what regards the “in-sample” fit to observed TSIR, in large measure due to having more degrees of freedom (weights and factors) than N&S (whose weights are parameterized). On the other hand, the N&S approach benefits from generating non-negative forward rates across all maturity spectrum, while we do not know in what measure PCA simulations are exposed to this possibility, given that they are not necessarily smooth as are those simulated by the N&S methodology.

With the exception made for skewness and kurtosis, the simulations presented in this work successfully describe the historical distribution of interest rates, spreads and curvatures. This applies to whatever time horizon we choose to make the simulations. The best improvements that could be brought to the models are those that could replicate the historically observed non-normality of interest rates.

Finally, and considering the remark that has just been made, the N&S methodology is reliable to simulate TSIR.

10. Further Research

As said before, improvements can be made to this study by analyzing the additional explaining power of Nelson and Siegel extensions, in the research line of Pooter (2007). The third factor explains less than the second and a fourth factor would explain less than the third. Having in mind the PCA results, we should expect a fourth factor to have little incremental explaining power. On the other hand, problems and difficulties in the estimation of models with more parameters are extensively documented in the literature: potential non-identification issues, non-convergence and dependence on initial values.

Another improvement would be to make the decay parameter time-dependent. According to Pooter (2007) results, the improvement in the fitting of the model when we allow the decay parameter to vary is less than when we add a fourth factor related to the curvature.

Finally, we should consider in the models the interactions between the evolution of TSIR and fundamental macro variables, such as inflation expectations and the output gap. We know that the level of TSIR is related with inflation (as proxy of inflation expectations) and the slope is linked to the cyclical dynamics of the economy given by the proportion of installed productive capacity that is actually used, for instance.

In the case of the third factor, there are no suggestions of related macro variables in current literature, although given the pace of scientific production in this area more results are expected soon.

11. Bibliography

- Alexander, C. (2001), “Market Models: A Guide to Financial Data Analysis”, John Wiley & Sons
- Ang, A. and Piazzesi, M. (2003), “A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables”, Journal of Monetary Economics, Vol. 50, 745-787
- BIS (2005), “Zero-coupon yield curves: technical documentation”, Bank for International Settlements, BIS Papers, No. 25
- Bliss, R. R. (1996), “Testing the Term Structure Estimation Methods”, Federal Reserve Bank of Atlanta Working Paper 96-12a
- Chou, Shuo-hong (2006), “The Predictability in the Shape of the Term Structure of Interest Rates: Empirical Study in Taiwan Government Bonds”, E-Thesis, School of Finance, National Taiwan University of Science and Technology (http://pc01.lib.ntust.edu.tw/ETD-db)
- Diebold, F. and Li, C. (2006), “Forecasting the term structure of government bond yields”, Journal of Econometrics, vol. 130, No. 2, 337-364
- Diebold, F., Piazzesi, M. and Rudebusch, G. (2005), “Modeling Bond Yields in Finance and Macroeconomics”, Working Paper
- Diebold, F., Rudebusch, G. and Aruoba, B. (2005), “The macroeconomy and the yield curve: a dynamic latent factor approach”, Journal of Econometrics, vol. 131, 309-338
- Fabozzi, F., Martellini, L. and Priaulet, P. (2005), “Predictability in the Shape of Term Structure of Interest Rates”, The Journal of Fixed Income, vol. 15, 40-53
- Galitz, L. (1995), Financial Engineering - Tools and Techniques to Manage Financial Risk, Revised Edition, Great Britain, FT Pitman Publishing

- Geyer, A. and Mader, R. (1999), “Estimation of the Term Structure of Interest Rates; a Parametric Approach”, Working Paper 37, Oesterreichische Nationalbank
- Litterman, R. and Scheinkman, J. (1991), “Common Factors Affecting Bond Returns”, Journal of Fixed Income, 1, 54-61.
- Luna, F. E. (2006), “Aplicação da metodologia de componentes principais na análise da estrutura a termo de taxa de juros brasileira e no cálculo do valor em risco”, Texto para Discussão,n.º 1146, IPEA
- Machado, J. F. , Marques, C. R., Neves, P. D. and Afonso, A., G. (2001), “Using the First Principal Component as a Core Inflation Indicator”, Banco de Portugal, Economic Research Department WP 9-01, September
- Nelson, C. and Siegel, A. (1987), Parsimonious Modeling of Yield Curve, Journal of Business, vol. 60, 473-489
- Pooter, M. (2007), “Examining the Nelson-Siegel class of term structure models”, Tinbergen Institute, Erasmus University Rotterdam Working Paper
- Rudebusch and Wu (2004), “A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy”, Working Paper, Federal Reserve Bank of San Francisco
- Stock, J. H. and Watson, M. W. (1988), “Variable Trends in Economic Time Series”, The Journal of Economic Perspectives, Vol. 2, No. 3, pp. 147-174
- Svensson, L. E. (1994), “Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994”, Centre for Economic Policy Research Discussion Paper 1051